SEQUENCE LISTING

GENERAL INFORMATION:

(i) APPLICANTS: Chen, Yao-Tseng; Stockert, Elisabeth;
Chen, Yachi; Garin-Chesa, Pilar; Rettig, Wolfgang J.;
van der Bruggen, Pierre; Boon-Falleur, Thierry;
Old, Lloyd J.

(ii) TITLE OF INVENTION: MONOCLONAL ANTIBODIES WHICH BIND TO
TUMOR REJECTION ANTIGEN PRECURSOR MAGE-1, RECOMBINANT MAGE-1,
AND MAGE-1 DERIVED IMMUNOGENIC PEPTIDES

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Felfe & Lynch
(B) STREET: 805 Third Avenue
(C) CITY: New York City
(D) STATE: New York
(F) ZIP: 10022

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
(B) COMPUTER: IBM
(C) OPERATING SYSTEM: PC-DOS
(D) SOFTWARE: Wordperfect

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 08,190,411
(B) FILING DATE: 01-FEBRUARY-1994
(C) CLASSIFICATION: 435

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 037,230
(B) FILING DATE: 26-MARCH-1993

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/US92/04354
(B) FILING DATE: 22-MAY-1992

(viii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 07/807,043
(B) FILING DATE: 12-DECEMBER-1991

(ix) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 07/764,364
(B) FILING DATE: 23-SEPTEMBER-1991

(x) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 07/728,838
(b) FILING DATE: 9-JULY-1991

(xi) PRIOR APPLICATION DATA:
    (A) APPLICATION NUMBER: 07/705,702
    (B) FILING DATE: 23-MAY-1991

(xii) ATTORNEY/AGENT INFORMATION:
    (A) NAME: Hanson, Norman D.
    (B) REGISTRATION NUMBER: 30,946
    (C) REFERENCE/DOCKET NUMBER: LUD 354

(xiii) TELECOMMUNICATION INFORMATION:
    (A) TELEPHONE: (212) 688-9200
    (B) TELEFAX: (212) 838-3884

(2) INFORMATION FOR SEQUENCE ID NO: 1:
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: ~~5724~~ 5627 base pairs
        (B) TYPE: nucleic acid
        (D) TOPOLOGY: linear
    (ii) MOLECULE TYPE: genomic DNA
    (ix) FEATURE:
        (A) NAME/KEY: MAGE-1 gene
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

10/19/98

```
CCCGGGGCAC CACTGGCATC CCTCCCCCTA CCACCCCCAA TCCCTCCCTT        50
TACGCCACCC ATCCAAACAT CTTCACGCTC ACCCCCAGCC CAAGCCAGGC       100
AGAATCCGGT TCCACCCCTG CTCTCAACCC AGGGAAGCCC AGGTGCCCAG       150
ATGTGACGCC ACTGACTTGA GCATTAGTGG TTAGAGAGAA GCGAGGTTTT       200
CGGTCTGAGG GGCGGCTTGA GATCGGTGGA GGGAAGCGGG CCCAGCTCTG       250
TAAGGAGGCA AGGTGACATG CTGAGGGAGG ACTGAGGACC CACTTACCCC       300
AGATAGAGGA CCCCAAATAA TCCCTTCATG CCAGTCCTGG ACCATCTGGT       350
GGTGGACTTC TCAGGCTGGG CCACCCCCAG CCCCCTTGCT GCTTAAACCA       400
CTGGGGACTC GAAGTCAGAG CTCCGTGTGA TCAGGGAAGG GCTGCTTAGG       450
AGAGGGCAGC GTCCAGGCTC TGCCAGACAT CATGCTCAGG ATTCTCAAGG       500
AGGGCTGAGG GTCCCTAAGA CCCCACTCCC GTGACCCAAC CCCCACTCCA       550
ATGCTCACTC CCGTGACCCA ACCCCCTCTT CATTGTCATT CCAACCCCCA       600
CCCCACATCC CCCACCCCAT CCCTCAACCC TGATGCCCAT CCGCCCAGCC       650
ATTCCACCCT CACCCCCACC CCCACCCCCA CGCCCACTCC CACCCCCACC       700
CAGGCAGGAT CCGGTTCCCG CCAGGAAACA TCCGGGTGCC CGGATGTGAC       750
GCCACTGACT TGCGCATTGT GGGGCAGAGA GAAGCGAGGT TTCCATTCTG       800
AGGGACGGCG TAGAGTTCGG CCGAAGGAAC CTGACCCAGG CTCTGTGAGG       850
AGGCAAGGTG AGAGGCTGAG GGAGGACTGA GGACCCCGCC ACTCCAAATA       900
GAGAGCCCCA AATATTCCAG CCCCGCCCTT GCTGCCAGCC CTGGCCCACC       950
CGCGGGAAGA CGTCTCAGCC TGGGCTGCCC CCAGACCCCT GCTCCAAAAG      1000
CCTTGAGAGA CACCAGGTTC TTCTCCCCAA GCTCTGGAAT CAGAGGTTGC      1050
TGTGACCAGG GCAGGACTGG TTAGGAGAGG GCAGGGCACA GGCTCTGCCA      1100
GGCATCAAGA TCAGCACCCA AGAGGGAGGG CTGTGGGCCC CCAAGACTGC      1150
ACTCCAATCC CCACTCCCAC CCCATTCGCA TTCCCATTCC CCACCCAACC      1200
CCCATCTCCT CAGCTACACC TCCACCCCCA TCCCTACTCC TACTCCGTCA      1250
CCTGACCACC ACCCTCCAGC CCCAGCACCA GCCCCAACCC TTCTGCCACC      1300
```

```
TCACCCTCAC TGCCCCCAAC CCCACCCTCA TCTCTCTCAT GTGCCCCACT        1350
CCCATCGCCT CCCCCATTCT GGCAGAATCC GGTTTGCCCC TGCTCTCAAC        1400
CCAGGGAAGC CCTGGTAGGC CCGATGTGAA ACCACTGACT TGAACCTCAC        1450
AGATCTGAGA GAAGCCAGGT TCATTTAATG GTTCTGAGGG GCGGCTTGAG        1500
ATCCACTGAG GGGAGTGGTT TTAGGCTCTG TGAGGAGGCA AGGTGAGATG        1550
CTGAGGGAGG ACTGAGGAGG CACACACCCC AGGTAGATGG CCCCAAAATG        1600
ATCCAGTACC ACCCCTGCTG CCAGCCCTGG ACCACCCGGC CAGGACAGAT        1650
GTCTCAGCTG GACCACCCCC CGTCCCGTCC CACTGCCACT TAACCCACAG        1700
GGCAATCTGT AGTCATAGCT TATGTGACCG GGGCAGGGTT GGTCAGGAGA        1750
GGCAGGGCCC AGGCATCAAG GTCCAGCATC CGCCCGGCAT TAGGGTCAGG        1800
ACCCTGGGAG GGAACTGAGG GTTCCCCACC CACACCTGTC TCCTCATCTC        1850
CACCGCCACC CCACTCACAT TCCCATACCT ACCCCTACC CCCAACCTCA         1900
TCTTGTCAGA ATCCCTGCTG TCAACCCACG GAAGCCACGG GAATGGCGGC        1950
CAGGCACTCG GATCTTGACG TCCCCATCCA GGGTCTGATG GAGGGAAGGG        2000
GCTTGAACAG GGCCTCAGGG GAGCAGAGGG AGGGCCCTAC TGCGAGATGA        2050
GGGAGGCCTC AGAGGACCCA GCACCCTAGG ACACCGCACC CCTGTCTGAG        2100
ACTGAGGCTG CCACTTCTGG CCTCAAGAAT CAGAACGATG GGGACTCAGA        2150
TTGCATGGGG GTGGGACCCA GGCCTGCAAG GCTTACGCGG AGGAAGAGGA        2200
GGGAGGACTC AGGGGACCTT GGAATCCAGA TCAGTGTGGA CCTCGGCCCT        2250
GAGAGGTCCA GGGCACGGTG GCCACATATG GCCCATATTT CCTGCATCTT        2300
TGAGGTGACA GGACAGAGCT GTGGTCTGAG AAGTGGGGCC TCAGGTCAAC        2350
AGAGGGAGGA GTTCCAGGAT CCATATGGCC CAAGATGTGC CCCCTTCATG        2400
AGGACTGGGG ATATCCCCGG CTCAGAAAGA AGGGACTCCA CACAGTCTGG        2450
CTGTCCCCTT TTAGTAGCTC TAGGGGGACC AGATCAGGGA TGGCGGTATG        2500
TTCCATTCTC ACTTGTACCA CAGGCAGGAA GTTGGGGGGC CCTCAGGGAG        2550
ATGGGGTCTT GGGGTAAAGG GGGGATGTCT ACTCATGTCA GGGAATTGGG        2600
GGTTGAGGAA GCACAGGCGC TGGCAGGAAT AAAGATGAGT GAGACAGACA        2650
AGGCTATTGG AATCCACACC CCAGAACCAA AGGGGTCAGC CCTGGACACC        2700
TCACCCAGGA TGTGGCTTCT TTTTCACTCC TGTTTCCAGA TCTGGGGCAG        2750
GTGAGGACCT CATTCTCAGA GGGTGACTCA GGTCAACGTA GGGACCCCCA        2800
TCTGGTCTAA AGACAGAGCG GTCCCAGGAT CTGCCATGCG TTCGGGTGAG        2850
GAACATGAGG GAGGACTGAG GGTACCCCAG GACCAGAACA CTGAGGGAGA        2900
CTGCACAGAA ATCAGCCCTG CCCCTGCTGT CACCCCAGAG AGCATGGGCT        2950
GGGCCGTCTG CCGAGGTCCT TCCGTTATCC TGGGATCATT GATGTCAGGG        3000
ACGGGGAGGC CTTGGTCTGA GAAGGCTGCG CTCAGGTCAG TAGAGGGAGC        3050
GTCCCAGGCC CTGCCAGGAG TCAAGGTGAG GACCAAGCGG GCACCTCACC   ~~3150~~ 3100
CAGGACACAT TAATTCCAAT GAATTTTGAT ATCTCTTGCT GCCCTTCCCC   ~~3200~~ 3150
AAGGACCTAG GCACGTGTGG CCAGATGTTT GTCCCCTCCT GTCCTTCCAT   ~~3250~~ 3200
TCCTTATCAT GGATGTGAAC TCTTGATTTG GATTTCTCAG ACCAGCAAAA   ~~3300~~ 3250
GGGCAGGATC CAGGCCCTGC CAGGAAAAAT ATAAGGGCCC TGCGTGAGAA   ~~3350~~ 3300
CAGAGGGGGT CATCCACTGC ATGAGAGTGG GGATGTCACA GAGTCCAGCC   ~~3400~~ 3350
CACCCTCCTG GTAGCACTGA GAAGCCAGGG CTGTGCTTGC GGTCTGCACC   ~~3450~~ 3400
CTGAGGGCCC GTGGATTCCT CTTCCTGGAG CTCCAGGAAC CAGGCAGTGA   ~~3500~~ 3450
GGCCTTGGTC TGAGACAGTA TCCTCAGGTC ACAGAGCAGA GGATGCACAG   ~~3550~~ 3500
GGTGTGCCAG CAGTGAATGT TTGCCCTGAA TGCACACCAA GGGCCCCACC   ~~3600~~ 3550
TGCCACAGGA CACATAGGAC TCCACAGAGT CTGGCCTCAC CTCCCTACTG   ~~3650~~ 3600
TCAGTCCTGT AGAATCGACC TCTGCTGGCC GGCTGTACCC TGAGTACCCT   ~~3700~~ 3650
CTCACTTCCT CCTTCAGGTT TTCAGGGGAC AGGCCAACCC AGAGGACAGG   ~~3750~~ 3700
ATTCCCTGGA GGCCACAGAG GAGCACCAAG GAGAAGATCT GTAAGTAGGC   ~~3800~~ 3750
CTTTGTTAGA GTCTCCAAGG TTCAGTTCTC AGCTGAGGCC TCTCACACAC   ~~3850~~ 3800
TCCCTCTCTC CCCAGGCCTG TGGGTCTTCA TTGCCCAGCT CCTGCCCACA   ~~3900~~ 3850
CTCCTGCCTG CTGCCCTGAC GAGAGTCATC                          ~~3930~~ 3880
ATG TCT CTT GAG CAG AGG AGT CTG CAC TGC AAG CCT GAG GAA   ~~3972~~ 3922
GCC CTT GAG GCC CAA CAA GAG GCC CTG GGC CTG GTG TGT GTG   ~~4014~~ 3964
```

```
CAG GCT GCC ACC TCC TCC TCC TCT CCT CTG GTC CTG GGC ACC        4056 4006
CTG GAG GAG GTG CCC ACT GCT GGG TCA ACA GAT CCT CCC CAG        4098 4048
AGT CCT CAG GGA GCC TCC GCC TTT CCC ACT ACC ATC AAC TTC        4140 4090
ACT CGA CAG AGG CAA CCC AGT GAG GGT TCC AGC AGC CGT GAA        4182 4132
GAG GAG GGG CCA AGC ACC TCT TGT ATC CTG GAG TCC TTG TTC        4224 4174
CGA GCA GTA ATC ACT AAG AAG GTG GCT GAT TTG GTT GGT TTT        4266 4216
CTG CTC CTC AAA TAT CGA GCC AGG GAG CCA GTC ACA AAG GCA        4308 4258
GAA ATG CTG GAG AGT GTC ATC AAA AAT TAC AAG CAC TGT TTT        4350 4300
CCT GAG ATC TTC GGC AAA GCC TCT GAG TCC TTG CAG CTG GTC        4392 4342
TTT GGC ATT GAC GTG AAG GAA GCA GAC CCC ACC GGC CAC TCC        4434 4384
TAT GTC CTT GTC ACC TGC CTA GGT CTC TCC TAT GAT GGC CTG        4476 4426
CTG GGT GAT AAT CAG ATC ATG CCC AAG ACA GGC TTC CTG ATA        4518 4468
ATT GTC CTG GTC ATG ATT GCA ATG GAG GGC GGC CAT GCT CCT        4560 4510
GAG GAG GAA ATC TGG GAG GAG CTG AGT GTG ATG GAG GTG TAT        4602 4552
GAT GGG AGG GAG CAC AGT GCC TAT GGG GAG CCC AGG AAG CTG        4644 4594
CTC ACC CAA GAT TTG GTG CAG GAA AAG TAC CTG GAG TAC GGC        4686 4636
AGG TGC CGG ACA GTG ATC CCG CAC GCT ATG AGT TCC TGT GGG        4728 4678
GTC CAA GGG CCC TCG CTG AAA CCA GCT ATG TGA                    4761 4711
AAGTCCTTGA GTATGTGATC AAGGTCAGTG CAAGAGTTC                      4800 4750
GCTTTTTCTT CCCATCCCTG CGTGAAGCAG CTTTGAGAGA GGAGGAAGAG          4850 4800
GGAGTCTGAG CATGAGTTGC AGCCAAGGCC AGTGGGAGGG GGACTGGGCC          4900 4850
AGTGCACCTT CCAGGGCCGC GTCCAGCAGC TTCCCCTGCC TCGTGTGACA          4950 4900
TGAGGCCCAT TCTTCACTCT GAAGAGAGCG GTCAGTGTTC TCAGTAGTAG          5000 4950
GTTTCTGTTC TATTGGGTGA CTTGGAGATT TATCTTTGTT CTCTTTTGGA          5050 5000
ATTGTTCAAA TGTTTTTTTT TAAGGGATGG TTGAATGAAC TTCAGCATCC          5100 5050
AAGTTTATGA ATGACAGCAG TCACACAGTT CTGTGTATAT AGTTTAAGGG          5150 5100
TAAGAGTCTT GTGTTTTATT CAGATTGGGA AATCCATTCT ATTTTGTGAA          5200 5150
TTGGGATAAT AACAGCAGTG GAATAAGTAC TTAGAAATGT GAAAAATGAG          5250 5200
CAGTAAAATA GATGAGATAA AGAACTAAAG AAATTAAGAG ATAGTCAATT          5300 5250
CTTGCCTTAT ACCTCAGTCT ATTCTGTAAA ATTTTTAAAG ATATATGCAT          5350 5300
ACCTGGATTT CCTTGGCTTC TTTGAGAATG TAAGAGAAAT TAAATCTGAA          5400 5350
TAAAGAATTC TTCCTGTTCA CTGGCTCTTT TCTTCTCCAT GCACTGAGCA          5450 5400
TCTGCTTTTT GGAAGGCCCT GGGTTAGTAG TGGAGATGCT AAGGTAAGCC          5500 5450
AGACTCATAC CCACCCATAG GGTCGTAGAG TCTAGGAGCT GCAGTCACGT          5550 5500
AATCGAGGTG GCAAGATGTC CTCTAAAGAT GTAGGGAAAA GTGAGAGAGG          5600 5550
GGTGAGGGTG TGGGGCTCCG GGTGAGAGTG GTGGAGTGTC AATGCCCTGA          5650 5600
GCTGGGGCAT TTTGGGCTTT GGGAAACTGC AGTTCCTTCT GGGGGAGCTG          5700 5650
ATTGTAATGA TCTTGGGTGG ATCC                                      5724 5674
```

<div style="margin-left">10/19/98 PAD</div>

(2)   INFORMATION FOR SEQUENCE ID NO: 2:
       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 14 amino acids
           (B) TYPE: amino acid
           (D) TOPOLOGY: linear
       (ii) MOLECULE TYPE: protein
       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Ile Asn Phe Thr Arg Gln Arg Gln Pro Ser Glu Gly Ser Ser
                    5   10

(2)   INFORMATION FOR SEQUENCE ID NO: 3:
      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 12 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear
      (ii) MOLECULE TYPE: protein
      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Leu Phe Arg Ala Val Ile Thr Lys Lys Val Ala Asp
                  5   10

(2)   INFORMATION FOR SEQUENCE ID NO: 4:
      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 12 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear
      (ii) MOLECULE TYPE: protein
      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Asp Val Lys Glu Ala Asp Pro Thr Gly His Ser Tyr
                  5   10

(1) GENERAL INFORMATION:

JU 253.6-JEL/NDH

(iii) NUMBER OF SEQUENCES: 4

(2) INFORMATION FOR SEQUENCE ID NO: 1:
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: ~5724 base pairs
        (B) TYPE: nucleic acid
        (D) TOPOLOGY: linear
    (ii) MOLECULE TYPE: genomic DNA
    (ix) FEATURE:
        (A) NAME/KEY: MAGE-1 gene
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
CCCGGGGCAC CACTGGCATC CCTCCCCCTA CCACCCCCAA TCCCTCCCTT        50
TACGCCACCC ATCCAAACAT CTTCACGCTC ACCCCCAGCC CAAGCCAGGC       100
AGAATCCGGT TCCACCCCTG CTCTCAACCC AGGGAAGCCC AGGTGCCCAG       150
ATGTGACGCC ACTGACTTGA GCATTAGTGG TTAGAGAGAA GCGAGGTTTT       200
CGGTCTGAGG GGCGGCTTGA GATCGGTGGA GGGAAGCGGG CCCAGCTCTG       250
TAAGGAGGCA AGGTGACATG CTGAGGGAGG ACTGAGGACC CACTTACCCC       300
AGATAGAGGA CCCCAAATAA TCCCTTCATG CCAGTCCTGG ACCATCTGGT       350
GGTGGACTTC TCAGGCTGGG CCACCCCCAG CCCCCTTGCT GCTTAAACCA       400
CTGGGGACTC GAAGTCAGAG CTCCGTGTGA TCAGGGAAGG GCTGCTTAGG       450
AGAGGGCAGC GTCCAGGCTC TGCCAGACAT CATGCTCAGG ATTCTCAAGG       500
AGGGCTGAGG GTCCCTAAGA CCCCACTCCC GTGACCCAAC CCCCACTCCA       550
ATGCTCACTC CCGTGACCCA ACCCCCTCTT CATTGTCATT CCAACCCCCA       600
CCCCACATCC CCCACCCCAT CCCTCAACCC TGATGCCCAT CCGCCCAGCC       650
ATTCCACCCT CACCCCCACC CCCACCCCCA CGCCCACTCC CACCCCCACC       700
CAGGCAGGAT CCGGTTCCCG CCAGGAAACA TCCGGGTGCC CGGATGTGAC       750
GCCACTGACT TGCGCATTGT GGGGCAGAGA GAAGCGAGGT TTCCATTCTG       800
AGGGACGGCG TAGAGTTCGG CCGAAGGAAC CTGACCCAGG CTCTGTGAGG       850
AGGCAAGGTG AGAGGCTGAG GGAGGACTGA GGACCCCGCC ACTCCAAATA       900
GAGAGCCCCA AATATTCCAG CCCCGCCCTT GCTGCCAGCC CTGGCCCACC       950
CGCGGGAAGA CGTCTCAGCC TGGGCTGCCC CCAGACCCCT GCTCCAAAAG      1000
CCTTGAGAGA CACCAGGTTC TTCTCCCCAA GCTCTGGAAT CAGAGGTTGC      1050
TGTGACCAGG GCAGGACTGG TTAGGAGAGG GCAGGGCACA GGCTCTGCCA      1100
GGCATCAAGA TCAGCACCCA AGAGGGAGGG CTGTGGGCCC CCAAGACTGC      1150
ACTCCAATCC CCACTCCCAC CCCATTCGCA TTCCCATTCC CCACCCAACC      1200
CCCATCTCCT CAGCTACACC TCCACCCCCA TCCCTACTCC TACTCCGTCA      1250
CCTGACCACC ACCCTCCAGC CCCAGCACCA GCCCCAACCC TTCTGCCACC      1300
TCACCCTCAC TGCCCCCAAC CCCACCCTCA TCTCTCTCAT GTGCCCCACT      1350
CCCATCGCCT CCCCCATTCT GGCAGAATCC GGTTTGCCCC TGCTCTCAAC      1400
CCAGGGAAGC CCTGGTAGGC CCGATGTGAA ACCACTGACT TGAACCTCAC      1450
AGATCTGAGA GAAGCCAGGT TCATTAATG GTTCTGAGGG GCGGCTTGAG       1500
ATCCACTGAG GGGAGTGGTT TTAGGCTCTG TGAGGAGGCA AGGTGAGATG      1550
CTGAGGGAGG ACTGAGGAGG CACACACCCC AGGTAGATGG CCCCAAAATG      1600
ATCCAGTACC ACCCCTGCTG CCAGCCCTGG ACCACCCGGC CAGGACAGAT      1650
GTCTCAGCTG GACCACCCCC CGTCCCGTCC CACTGCCACT TAACCCACAG      1700
GGCAATCTGT AGTCATAGCT TATGTGACCG GGGCAGGGTT GGTCAGGAGA      1750
GGCAGGGCCC AGGCATCAAG GTCCAGCATC CGCCCGGCAT TAGGGTCAGG      1800
ACCCTGGGAG GGAACTGAGG GTTCCCCACC CACACCTGTC TCCTCATCTC      1850
CACCGCCACC CCACTCACAT TCCCATACCT ACCCCCTACC CCCAACCTCA      1900
TCTTGTCAGA ATCCCTGCTG TCAACCCACG GAAGCCACGG GAATGGCGGC      1950
```

```
CAGGCACTCG GATCTTGACG TCCCCATCCA GGGTCTGATG GAGGGAAGGG    2000
GCTTGAACAG GGCCTCAGGG GAGCAGAGGG AGGGCCCTAC TGCGAGATGA    2050
GGGAGGCCTC AGAGGACCCA GCACCCTAGG ACACCGCACC CCTGTCTGAG    2100
ACTGAGGCTG CCACTTCTGG CCTCAAGAAT CAGAACGATG GGGACTCAGA    2150
TTGCATGGGG GTGGGACCCA GGCCTGCAAG GCTTACGCGG AGGAAGAGGA    2200
GGGAGGACTC AGGGGACCTT GGAATCCAGA TCAGTGTGGA CCTCGGCCCT    2250
GAGAGGTCCA GGGCACGGTG GCCACATATG GCCCATATTT CCTGCATCTT    2300
TGAGGTGACA GGACAGAGCT GTGGTCTGAG AAGTGGGGCC TCAGGTCAAC    2350
AGAGGGAGGA GTTCCAGGAT CCATATGGCC CAAGATGTGC CCCCTTCATG    2400
AGGACTGGGG ATATCCCCGG CTCAGAAAGA AGGGACTCCA CACAGTCTGG    2450
CTGTCCCCTT TTAGTAGCTC TAGGGGACC AGATCAGGGA TGGCGGTATG    2500
TTCCATTCTC ACTTGTACCA CAGGCAGGAA GTTGGGGGGC CCTCAGGGAG    2550
ATGGGGTCTT GGGGTAAAGG GGGGATGTCT ACTCATGTCA GGGAATTGGG    2600
GGTTGAGGAA GCACAGGCGC TGGCAGGAAT AAAGATGAGT GAGACAGACA    2650
AGGCTATTGG AATCCACACC CCAGAACCAA AGGGGTCAGC CCTGGACACC    2700
TCACCCAGGA TGTGGCTTCT TTTTCACTCC TGTTTCCAGA TCTGGGGCAG    2750
GTGAGGACCT CATTCTCAGA GGGTGACTCA GGTCAACGTA GGGACCCCCA    2800
TCTGGTCTAA AGACAGAGCG GTCCCAGGAT CTGCCATGCG TTCGGGTGAG    2850
GAACATGAGG GAGGACTGAG GGTACCCCAG GACCAGAACA CTGAGGGAGA    2900
CTGCACAGAA ATCAGCCCTG CCCCTGCTGT CACCCCAGAG AGCATGGGCT    2950
GGGCCGTCTG CCGAGGTCCT TCCGTTATCC TGGGATCATT GATGTCAGGG    3000
ACGGGGAGGC CTTGGTCTGA GAAGGCTGCG CTCAGGTCAG TAGAGGGAGC    3050
GTCCCAGGCC CTGCCAGGAG TCAAGGTGAG GACCAAGCGG GCACCTCACC    3̶1̶5̶0̶ 3100
CAGGACACAT TAATTCCAAT GAATTTTGAT ATCTCTTGCT GCCCTTCCCC    3̶2̶0̶0̶ 3150
AAGGACCTAG GCACGTGTGG CCAGATGTTT GTCCCCTCCT GTCCTTCCAT    3̶2̶5̶0̶ 3200
TCCTTATCAT GGATGTGAAC TCTTGATTTG GATTTCTCAG ACCAGCAAAA    3̶3̶0̶0̶ 3250
GGGCAGGATC CAGGCCCTGC CAGGAAAAAT ATAAGGGCCC TGCGTGAGAA    3̶3̶5̶0̶ 3300
CAGAGGGGGT CATCCACTGC ATGAGAGTGG GGATGTCACA GAGTCCAGCC    3̶4̶0̶0̶ 3350
CACCTCCTG GTAGCACTGA GAAGCCAGGG CTGTGCTTGC GGTCTGCACC    3̶4̶5̶0̶ 3400
CTGAGGGCCC GTGGATTCCT CTTCCTGGAG CTCCAGGAAC CAGGCAGTGA    3̶5̶0̶0̶ 3450
GGCCTTGGTC TGAGACAGTA TCCTCAGGTC ACAGAGCAGA GGATGCACAG    3̶5̶5̶0̶ 3500
GGTGTGCCAG CAGTGAATGT TTGCCCTGAA TGCACACCAA GGGCCCCACC    3̶6̶0̶0̶ 3550
TGCCACAGGA CACATAGGAC TCCACAGAGT CTGGCCTCAC CTCCCTACTG    3̶6̶5̶0̶ 3600
TCAGTCCTGT AGAATCGACC TCTGCTGGCC GGCTGTACCG TGAGTACCCT    3̶7̶0̶0̶ 3650
CTCACTTCCT CCTTCAGGTT TTCAGGGGAC AGGCCAACCC AGAGGACAGG    3̶7̶5̶0̶ 3700
ATTCCCTGGA GGCCACAGAG GAGCACCAAG GAGAAGATCT GTAAGTAGGC    3̶8̶0̶0̶ 3750
CTTTGTTAGA GTCTCCAAGG TTCAGTTCTC AGCTGAGGCC TCTCACACAC    3̶8̶5̶0̶ 3800
TCCCTCTCTC CCCAGGCCTG TGGGTCTTCA TTGCCCAGCT CCTGCCCACA    3̶9̶0̶0̶ 3850
CTCCTGCCTG CTGCCCTGAC GAGAGTCATC                          3̶9̶3̶0̶ 3880
ATG TCT CTT GAG CAG AGG AGT CTG CAC TGC AAG CCT GAG GAA    3̶9̶7̶2̶ 3922
GCC CTT GAG GCC CAA CAA GAG GCC CTG GGC CTG GTG TGT GTG    4̶0̶1̶4̶ 3964
CAG GCT GCC ACC TCC TCC TCC TCT CCT CTG GTC CTG GGC ACC    4̶0̶5̶6̶ 4006
CTG GAG GAG GTG CCC ACT GCT GGG TCA ACA GAT CCT CCC CAG    4̶0̶9̶8̶ 4048
AGT CCT CAG GGA GCC TCC GCC TTT CCC ACT ACC ATC AAC TTC    4̶1̶4̶0̶ 4090
ACT CGA CAG AGG CAA CCC AGT GAG GGT TCC AGC AGC CGT GAA    4̶1̶8̶2̶ 4132
GAG GAG GGG CCA AGC ACC TCT TGT ATC CTG GAG TCC TTG TTC    4̶2̶2̶4̶ 4174
CGA GCA GTA ATC ACT AAG AAG GTG GCT GAT TTG GTT GGT TTT    4̶2̶6̶6̶ 4216
CTG CTC CTC AAA TAT CGA GCC AGG GAG CCA GTC ACA AAG GCA    4̶3̶0̶8̶ 4258
GAA ATG CTG GAG AGT GTC ATC AAA AAT TAC AAG CAC TGT TTT    4̶3̶5̶0̶ 4300
CCT GAG ATC TTC GGC AAA GCC TCT GAG TCC TTG CAG CTG GTC    4̶3̶9̶2̶ 4342
TTT GGC ATT GAC GTG AAG GAA GCA GAC CCC ACC GGC CAC TCC    4̶4̶3̶4̶ 4384
```

21

```
TAT GTC CTT GTC ACC TGC CTA GGT CTC TCC TAT GAT GGC CTG          4476  4426
CTG GGT GAT AAT CAG ATC ATG CCC AAG ACA GGC TTC CTG ATA          4518  4468
ATT GTC CTG GTC ATG ATT GCA ATG GAG GGC GGC CAT GCT CCT          4560  4510
GAG GAG GAA ATC TGG GAG GAG CTG AGT GTG ATG GAG GTG TAT          4602  4552
GAT GGG AGG GAG CAC AGT GCC TAT GGG GAG CCC AGG AAG CTG          4644  4594
CTC ACC CAA GAT TTG GTG CAG GAA AAG TAC CTG GAG TAC GGC          4686  4636
AGG TGC CGG ACA GTG ATC CCG CAC GCT ATG AGT TCC TGT GGG          4728  4678
GTC CAA GGG CCC TCG CTG AAA CCA GCT ATG TGA                      4761  4711
AAGTCCTTGA GTATGTGATC AAGGTCAGTG CAAGAGTTC                        4800  4750
GCTTTTTCTT CCCATCCCTG CGTGAAGCAG CTTTGAGAGA GGAGGAAGAG           4850  4800
GGAGTCTGAG CATGAGTTGC AGCCAAGGCC AGTGGGAGGG GGACTGGGCC           4900  4850
AGTGCACCTT CCAGGGCCGC GTCCAGCAGC TTCCCCTGCC TCGTGTGACA           4950  4900
TGAGGCCCAT TCTTCACTCT GAAGAGAGCG GTCAGTGTTC TCAGTAGTAG           5000  4950
GTTTCTGTTC TATTGGGTGA CTTGGAGATT TATCTTTGTT CTCTTTTGGA           5050  5000
ATTGTTCAAA TGTTTTTTTT TAAGGGATGG TTGAATGAAC TTCAGCATCC           5100  5050
AAGTTTATGA ATGACAGCAG TCACACAGTT CTGTGTATAT AGTTTAAGGG           5150  5100
TAAGAGTCTT GTGTTTATT CAGATTGGGA AATCCATTCT ATTTTGTGAA            5200  5150
TTGGGATAAT AACAGCAGTG GAATAAGTAC TTAGAAATGT GAAAAATGAG           5250  5200
CAGTAAAATA GATGAGATAA AGAACTAAAG AAATTAAGAG ATAGTCAATT           5300  5250
CTTGCCTTAT ACCTCAGTCT ATTCTGTAAA ATTTTTAAAG ATATATGCAT           5350  5300
ACCTGGATTT CCTTGGCTTC TTTGAGAATG TAAGAGAAAT TAAATCTGAA           5400  5350
TAAAGAATTC TTCCTGTTCA CTGGCTCTTT TCTTCTCCAT GCACTGAGCA           5450  5400
TCTGCTTTTT GGAAGGCCCT GGGTTAGTAG TGGAGATGCT AAGGTAAGCC           5500  5450
AGACTCATAC CCACCCATAG GGTCGTAGAG TCTAGGAGCT GCAGTCACGT           5550  5500
AATCGAGGTG GCAAGATGTC CTCTAAAGAT GTAGGGAAAA GTGAGAGAGG           5600  5550
GGTGAGGGTG TGGGGCTCCG GGTGAGAGTG GTGGAGTGTC AATGCCCTGA           5650  5600
GCTGGGGCAT TTTGGGCTTT GGGAAACTGC AGTTCCTTCT GGGGGAGCTG           5700  5650
ATTGTAATGA TCTTGGGTGG ATCC                                       5724  5674
```

22